Scheduling Packet Flows in Multi-Rate Wireless Local Area Networks

Field of the Invention

[01]   This invention relates generally to wireless local area networks, and more particularly to scheduling series of data packets for transmission in such networks.

Background of the Invention

[02]   A local area network (LAN) is a communications network that spans a relatively small area. Most LANs are confined to a single building or group of buildings. However, LANs can be connected to other networks, such as the Internet, using communication channels. LANs can locally connect terminal devices such as workstations, personal computers, printers, entertainment devices, sensors, and telephones. There are many different types of LANs, Ethernet being the most common.

[03]   A wireless local area network (WLAN) is a shared-medium communication network that transmits information over wireless channels between terminals within radio transmission range. IEEE 802.11 is a communication standard for WLANs, see O'Hara et al., "IEEE 802.11 Handbook, A Designer's Companion," IEEE Press, 1999, and "IEEE 802.11e/D5.0, Draft Supplement to Part 11:Wireless Medium Access Control (MAC) and physical layer (PHY)specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)." The medium access method of (MAC) of IEEE 802.11 is called a distributed coordination

function (DCF), which is based on carrier sense multiple access/collision avoidance (CSMA/CA) protocol.

[04]   IEEE 802.11 wireless networks can be configured in two different modes: ad hoc and infrastructure modes. In ad hoc mode, all wireless terminals within communication range can communicate directly with each other. The present invention can be applied to Wireless LAN operating both in infrastructure mode and ad hoc mode. The IEEE 802.11a standard specifies eight transmission bit rates at 6, 9, 12, 18, 24, 36, 48, and 54 Mbps, and the 802.11b standard specifies four rate at 1, 2, 5.5, and 11 Mbps. With such a physical-layer enhancement, a terminal can select a best transmission rate depending on a quality of the wireless channel measured by the signal-to-noise ratio (SNR) or packet error rate. When used properly, a multi-rate WLAN greatly improves throughput and effectively supports communication-intensive multimedia applications.

[05]   In the context of scheduling a series of packets between terminals, i.e., a 'packet flow', the multi-rate WLAN poses new challenges for designing network protocols. Packet scheduling, particularly fair queuing, provides packet-level quality of services (QoS) in terms of throughput and delay, to enable both delay-sensitive and throughput-sensitive applications. Wireless packet scheduling should also address location-dependent channel errors in wireless networks, and shields short-term error bursts from packet flows. Location-dependent errors are a particular concern when terminals are mobile.

[06]   However, prior art wireless fair packet scheduling typically assumes a single, fixed transmission rate for *all* terminals, and does not consider multiple rates for different terminals, where at least one terminal is transmitting at a different rate than all other terminals. As a result, fair queuing methods that are designed for single-rate networks suffer from significant throughput reduction in WLANs based on the 802.11a/b/g standards. In fact, even the prior art fairness notion may not be justified.

[07]   In wireless fair scheduling, packet flows, in the presence of channel errors, seek to approximate a QoS that is equivalent to idealistic, error-free channel conditions. To this end, transmissions of error-prone packet flows are deferred temporarily while error-free flows advance. This substantially improves the effective channel throughput because only flows that perceive error-free channels transmit packets at any given time.

[08]   Compared with error-free services, where all flows perceive error-free channels all of the time, error-prone flows may temporally lag behind and error-free flows lead ahead in short periods of time. However, these leading flows delay their future transmissions when the lagging flows perceive error-free channels in order to compensate for the time the lagging flows were deferred.

[09]   This way, both leading and lagging flows receive their contracted rates, i.e., throughput, over extended periods of time, and performance bounds in terms of throughput, delay and fairness are preserved over the extended periods of time in the presence of error-prone channels. In prior art design, both models of fair sharing among competing flows and

3

compensation for error-prone flows are defined with respect to throughput. That is, *'fairness'* means that over an extended period of time, all flows experience a *substantial equal throughput.*

[010]   Roughly speaking, each backlogged flow receives a 'fair' share of throughput, defined as *f* bytes per second. The compensation in the presence of channel errors is also based on throughput.

[011]   If an assigned transmission rate of a packet flow *f* is r_f, then the flow *f* receives services during a time period $[t, t + \Delta T]$ in proportion to the rate r_f, as given by

$$S_f(t, t + \Delta T) \approx \frac{r_f}{\sum_{i \in B} r_i} C \Delta T \; ,$$

where *C* is a channel capacity perceived by *all* flows, *B* denotes a set of flows that are back-logged due to experiencing an error prone channel, and ΔT is some predetermined time increment.

[012]   While such a model is adequate for single-rate WLANS, that model does not work for multi-rate WLANs because the flows have different rates and there is no single channel capacity *C*. Instead, the channel capacity varies over time depending on the rates assigned to the various terminals. Moreover, normalizing flow throughputs in a multi-rate network leads to significant inefficiencies and mitigates the gains offered by the multi-rate physical layer, because error-prone flows consume a disproportional amount of time and channel resources.

[013]   If each terminal in a WLAN can potentially assign different transmission rates at different times, i.e., at least one terminal is transmitting at a different rate than all other terminals, then the fundamental problem is to adapt the compensation models to accommodate such multi-rate options in wireless networks.

Summary of the Invention

[014]   The invention provides a multi-rate wireless fair scheduling (MWFS) method for data packets in a wireless local area network. The packet scheduler according to the invention utilizes the multi-rate capability by automatically selecting a best transmission rate to match a perceived channel quality.

[015]   A key notion of the MWFS method is to define *fairness* and *compensation* in terms of *temporal* shares and to ensure that backlogged packet flows receive *fair* shares in terms of transmission time over extended periods of time, e.g., seconds or minutes. As before, the term *fairness* means substantially equal, but in this case the fairness applies to *time*, instead of *throughput*, as in the prior art. The MWFS method achieves significant throughput gains by providing packet flows with dramatically different throughputs as governed by channel conditions.

[016]   The WMFS method can achieve significantly throughput gains in the order of 159% when compared to prior art fair throughput scheduling methods.

[017]   In particular, system and method schedule series of packets for transmission between terminals in a single wireless channel of a packet-switched local area network.

[018]   The local area network can operate in an ad hoc mode or in infrastructure mode. A transmission rate is assigned to each of the terminals. The terminals can be mobile.

[019]   Different terminals can be assigned different transmission rates, depending on a quality of the channel. That is, terminals transmitting via an error-free channel are assigned a higher transmission rate than terminals transmitting via an error-prone channel.

[020]   Then, the series of packets are scheduled for transmission between the terminals such that each terminal receives a substantially equal amount of transmission time over an extended period of time.

Brief Description of the Drawings

[021]   Figure 1A is a block diagram of a communications system including a multi-rate wireless fair scheduler according to the invention;

[022]   Figure 1B is a block diagram of the multi-rate wireless fair scheduler of Figure 1A;

[023]   Figure 2 is a graph comparing throughput in an error-free channel for prior art single rate fair scheduling and multi-rate wireless fair scheduling according to the invention;

[024]   Figure 3 is a table of packet delay statistics in an error-free channel for different packet source using the multi-rate wireless fair scheduling according to the invention;

[025]   Figure 4 is a graph comparing throughput in an error-prone channel for the prior art single rate fair scheduling and the multi-rate wireless fair scheduling according to the invention;

[026]   Figure 5 is a graph comparing packet delay for MMPP sources for the prior art single rate fair scheduling and the multi-rate wireless fair scheduling according to the invention; and

[027]   Figure 6 is a graph comparing packet delay for Poisson sources for the prior art single rate fair scheduling and the multi-rate wireless fair scheduling according to the invention.

Detailed Description of the Preferred Embodiment

[028]   As shown in Figure 1A, our invention provides a system and method 200 for scheduling series of data packets 101-102, i.e., '*packet flows*' in a packet-switched wireless local area network (WLAN) 100 via an access point 110. Particularly, different terminals 111-112 can communicate packets on a single physical transmission channel 105 at different assigned

transmission rates. That is, at least one terminal is transmitting packets at a different transmission rate than all other terminals.

[029]   We call our scheduling method 200 multi-rate wireless fair scheduling (MWFS). Our method can be used with WLANs based on the IEEE 802.11a/b/g/e standards. In essence, the method achieves temporal fairness in the presence of location-dependent channel errors. Location-dependent errors are a particular problem when the terminals are mobile.

[030]   As shown in Figure 1B, the scheduler 200 uses three models, an error-free service model 210, a compensation model 220, and a lead and lag model 230. The scheduler also uses a channel estimation module 250 and a channel access module 260.

[031]   The error-free service model 210 defines ideal packet flows that transmit at different rates over an error-free channel. The compensation model 220 compensates lagging packet flows at the expense of leading flows. The lead and lag model 230 determines which packet flows are leading or lagging their error-free service, and by how much. This addresses the problem of short-term error bursts, location-dependent channel errors due to mobile terminals, and channel contention in WLANs.

[032]   The pseudo-code for the scheduler 200 is given in the Appendix.

[033]   The channel estimation model 250 determines a quality of the channel. The channel access module 260 provides access to the physical channel 105.

[034]   The method according to the invention is backward compatible with prior art wireless fair scheduling (WFS) for single rate networks, e.g., see Lu et al., "Fair scheduling in wireless packet networks," IEEE/ACM Transactions on Networking, August, 1999. If all packet flows perceive identical transmission rates, then our MWFS method degenerates to the prior art WFS because all terminals will have the same throughput.

[035]   We apply our MWFS method to WLANs based on the IEEE 802.11b/a/g/e standards. There, terminals transmitting via error-free channels can use a higher assigned transmission rate than terminals using error-prone channels.

[036]   Our distinction between fairness based on *temporal* share and *throughput* share is important in multi-rate WLANs. Our MWFS method can provide packet flows with dramatically different throughputs as governed by channel conditions. However, all packet flows receive approximately identical shares of time over long periods of time, e.g., seconds or minutes.

[037]   Network Model

[038]   In the infrastructure mode of WLANs based on the IEEE 802.11b/a/g/e standard, the access point (AP) 110 coordinates the transmissions of the data packets 101-102 for the terminals 111-112 located within the transmission range of the AP 110. A single wireless channel 105 is shared for both uplink and a downlink communications between the terminals and the AP. Terminals communicate with each other via the AP. A

triplet {terminal, uplink/downlink flag, identification} designates each flow of packets.

[039]   Even though all terminals and the AP share the same channel, error rates can vary depending on the locations of the terminals with respect to the AP. Location-dependent errors are due to fading, interference, obstructions, and the like, particularly for mobile terminals.

[040]   A flow of packets is said to perceive an error-free channel when packets are transmitted successfully. Otherwise, if the packet transmission fails, then a flow is said to perceive an error-prone channel. Errors in the wireless channel typically occur in short bursts and are highly correlated in short periods of time, for example, the very small fraction of a second it takes to send a small number of packets, e.g., a dozen, but uncorrelated over extended periods of time when thousands of packets can be transmitted in a second.

[041]   We do not make any assumptions about the exact error model, though we do assume that the number of errors for a given time period T_i is bounded, i.e., a packet flow i receive no more than e_i packet errors in the given time period T_i.

[042]   Our underlying MAC protocol follows the 802.11 standard. In distributed coordination function (DCF) mode, each data packet transmission is preceded by a RTS/CTS handshake, and followed by an ACK packet. In PCF mode, each data packet transmission is preceded by a polling exchange between the AP and the terminal.

[043] Multi-Rate Wireless Fair Scheduling

[044]   Our MWFS method departs from the prior art throughput-oriented fairness and compensation models available for single-rate networks. The MWFS uses the notion of equal *temporal* shares. Therefore, the MWFS method can greatly improve overall throughput.

[045]   As additional benefit, the MWFS method is backward compatible with the prior art WFS for single-rate network. If all flows perceive identical transmission rates, then MWFS method degenerates to WFS method.

[046] Error-Free Service Model

[047]   In this section, we describe the basic scheduling model for the multi-rate wireless fair scheduling method according to the invention. For simplicity, we first consider an error-free service, i.e., a service where all of the terminals perceive an error-free channel. In subsequent sections, we describe services for error-prone channels.

[048]   Our model is an adaptation of the well known start-time fair queuing (SFQ) scheduling method, see Goyal et al., "Start-time Fair Queuing: A Scheduling Algorithm for Integrated Services Packet Switching Networks," IEEE/ACM Transactions on Networking, Vol. 5, No. 5, pp. 690-704, October 1997. In SFQ, each arriving packet is assigned two tags: a start tag and a finish tag.

[049]   Specifically, a packet with sequence number k of packet flow f arriving at time $A(t_k^f)$ is assigned two tags: a start tag S_k^f and a finish tag F_k^f, defined as follows:

$$S_k^f = \max\{V(A(t_k^f)), F_{k-1}^f\}$$

$$F_k^f = S_k^f + L_p / r_f, \tag{1}$$

where L_p denotes a size of the packet in bits, $V(.)$ is a virtual time for the start tag of the packet currently being served in the scheduler, and a base transmission rate of a flow f is r_f. Then, SFQ selects the packet flow with the minimum start tag and transmits its head-of-line (HoL) packet, i.e., the very next packet to be transmitted.

[050]   The invention adapts SFQ to the multi-rate WLAN. For a packet flow f, a current transmission rate at time t is $C_f(t)$. This current rate can be normalized with respect to the base transmission rate r_f. Therefore, the tagging as modified according to the invention is:

$$S_k^f = \max\{V(A(t_k^f)), F_{k-1}^f\}$$

$$F_k^f = S_k^f + L_p / (r_f \bullet C_f(t)), \tag{2}$$

where C_f is the current transmission rate, and the other variables are as defined for equation (1).

[051]   Therefore, both the start and the finish tags are normalized with respect to a current assigned transmission rate. This enables a high-rate packet flow, which perceives an error-free channel, to transmit at a higher transmission rate.

[052]   In addition, the method provides *temporal* fairness for each flow that is lagging. After tags are assigned to each packet, the scheduling allows the flow f_{min} with a *smallest* start tag to transmit first. The pseudo code of the above MWFS tagging is given in the function *enqueue*() of the Appendix.

[053]   Lead/Lag Model

[054]   The lead and lag model specifies how much compensation is needed for a lagging packet flow at the cost of a leading flow. In our MWFS method, the compensation is defined in terms of *time units* to ensure our fair *temporal* compensation. Each flow f has an associated credit counter $E_f(t)$ to indicate a status of a current flow, e.g., $E_f(t) > 0$ indicates that the flow f is leading, and a credit counter $E_f(t) < 0$ indicates that the flow is lagging. If $E_f(t) = 0$, then the flow is 'on' schedule.

[055]   When a flow is experiencing an error-prone channel, the flow relinquishes time units to other flows so that the flow becomes a lagging flow, and the other flows become leading flows. Later, when the lagging flow again experiences an error-free channel, the leading flows, in turn, relinquish time units back to the lagging flow to get it back on schedule.

[056]   Our design is based on two principles. First, we increment the credit counter for a leading flow *only* when the time units given up by a lagging flow g are used by a leading flow l. While at the same time, the credit counter of the lagging flow is decremented by the same amount. Second, the time units of the credit counter are expressed in terms of transmitted bytes, normalized with respect to the transmission rate of the

packet flow. That is, the incrementing can be expressed by $E_l(t) = E_l(t) + L_g/C_l(t)$, where L_p is a size of the packet transmitted by the leading packet flow, and $C_l(t)$ is a current transmission rate of transmission of the leading flow.

[057]   In essence, the lead and lag model specifies the temporal share, i.e., how many time units, a leading flow gives up in the future, and how many time units a lagging flow receives in compensation for having been deferred.

[058]   **Compensation Model**

[059]   Two additional issues need to be addressed. First, the scheduler needs to determine whether a leading flow is allowed to transmit a data packet or whether the current scheduling opportunity should be relinquished for compensation for a deferred lagging flow. Second, the scheduler needs to determine which lagging flow is allowed to transmit in the case the current scheduling opportunity has been relinquished by a leading flow and there are multiple lagging flows.

[060]   We use a graceful service relinquishing model for compensating lagging flows. Consider a leading packet flow i with a base transmission rate of r_f, a credit counter $E(i)$, and a *maximum* lead of $E_{max}(i)$. In other words, the flow with the largest lead is selected to relinquish time units to a lagging flow. This flow i is portioned hierarchically into a compensation flow i^c and a transmission flow i^t, with respective transmission rates of $r_i.E(i)/E_{max}(i)$, and $r_i.(1 - E(i)/E_{max}(i))$.

[061]   When the packet flow i is allocated a transmission opportunity, the flow i hierarchically schedules the opportunity among compensation flow i^c and transmission flow i^t. All time units belonging to the compensation flow i^c are relinquished to a lagging flow selected as described below. Note that as the lead decreases, the transmission rate for compensation flow i^c decreases linearly. This has the property of a graceful degradation of service for the leading flow.

[062]   After the leading flow gives up the transmission opportunity for its head-of-line packet, we need to select a lagging flow for transmission in the given time unit. If the packet size is L_i, and the transmission rate of the leading flow is $C_i(t)$, then the available time for compensation is

$\Delta T_i = L_i / C_i(t)$.

[063]   In a multi-rate network, the manner in which the lagging flow is selected to fill the time unit ΔT_i presents a new problem. In the WFS method for the single-rate network, a fixed time unit and identical packet size are assumed for all flows. Therefore, any lagging flow can be selected to transmit in the relinquished unit by the leading flow.

[064]   In the multi-rate case, this is not true anymore. In fact, the transmission of the head-of-line packet from the selected lagging flow can only be a portion of the available time unit ΔT_i, or can far exceed the length of the time unit ΔT_i. Moreover, the simple weighted round robin (WRR)

based compensation used by the WFS method is not applicable for the multi-rate network because the weights may not be whole integers.

[065] Our compensation model preserves the feature of graceful compensation among lagging flows even when there are different transmission rates and arbitrary sized packet.

[066] If the credit counter of a lagging flow i is $E_i(t)$, then we normalize the credit counter for all lagging flows according to

$g_i(t) = E_i(t)/\sum_{j \in F(t)} E_j(t)$, where $F(t)$ denotes all lagging flows at time t.

[067] Each lagging flow receives compensation, in terms of time units, in proportion to $g_i(t)$. This can be implemented with the following operations.

[068] The head-of-line packet from the lagging flow g is assigned a compensation tag $L_g/C_g(t) \cdot g_g(t)$, given that L_g is the size of the head-of-the-line packet, and $C_g(t)$ is the current transmission rate.

[069] Then, the lagging flow with the smallest compensation tag is selected to receive the relinquished compensation time unit. That is, our compensation model allocates compensation time units fairly, i.e., substantially equal, among all lagging flows.

[070]   Implementation

[071]   There are a number of implementation issues of our MWFS method for IEEE 802.11 MAC networks.

[072]   Channel State Estimation and Propagation

[073]   Because a receiving terminal has the most accurate information on the packet error rate, we estimate 250 a channel state, i.e., error-free or error-prone, in the receiving terminal. For a terminal that is not the intended recipient at a given instant in time, we can use the RTS/CTS handshake in the DCF mode or the polling message in the PCF mode to estimate channel states, see Bianchi, "Performance analysis of the IEEE 802.11 media access protocol for wireless LANs," Proceedings of ACM SIGCOMM '94, 1994. After estimating the channel state, the terminal can send the channel state to the AP as part of polling packets. The AP can then use the polling state to schedule packets appropriately.

[074]   Handling Uplink Flows

[075]   Our MWFS method works for both downlink and uplink flows. For uplink flows, the AP periodically polls each terminal to collect the following information: the size of the HoL packet, and the arrival time of the HoL packet if the packet flow lags again after an idle period.

[076]   The AP can acquire the transmission rate from each terminal. The details for this can be as described for ARF, Kamerman et al. "WaveLAN II:

A high-performance wireless LAN for the unlicensed band," Bell Labs Technical Journal, pp. 118-133, Summer 1997, or refined RBAR, see Holland et al., "A rate-adaptive MAC protocol for multi-hop wireless networks," Proceedings of ACM MOBICOM'01, 2001, or OAR, see Sadeghi et al., "OAR: A Multi-rate Media Access Protocol for Wireless Ad Hoc Networks," ACM Mobile Networking and Applications, 2003.

[077] Evaluation

[078] In this section, we describe simulation results to evaluate the performance of the MWFS method under various conditions. The reference prior art WFS method is the one described by Holland et al. That method utilizes a modified fluid fair algorithm that provides throughput fairness and throughput delay bounds for packet flows for both error-free and error-prone channels.

[079] The simulation setting is as follows. Four types of traffics are considered in the simulations shown in Figures 2-6, i.e., FTP sources, CBR sources, Poisson sources and Markov-modulated Poisson Process (MMPP) sources.

[080] The transmission rates are 2 Mbps, 5.5 Mbps and 11 Mbps according to the IEEE 802.11b standard. The packet size for each flow can vary. We use the well known two-state discrete Markov Chain to simulate channel errors, and the parameters p_g and p_e represent probabilities that the next time unit will experience an error-free channel, or an error-prone channel, respectively.

18

[081]   Therefore, the steady-state probabilities of the error-free channel P_G and the error-prone channel P_E are given by $P_G = p_g / (p_g + p_e)$ and $P_E = p_e / (p_g + p_e)$, respectively. We use a one-step prediction to estimate the channel state in the immediate future based on the current state of the channel. This exploits the reality that the channel errors are highly correlated over short periods of time, e.g., fractions of seconds. Each simulation run lasts for 100,000 time units, unless otherwise explicitly stated, and the results are averaged over 50 simulation runs.

[082]   Throughput in Error-Free Channels

[083]   In this set of simulations, we consider six FTP flows in an error-free channel to demonstrate that the throughput improvement of our MWFS method over the prior art WFS method. Each flow uses a 2 Mbps transmission rate as a comparison base. Then, we adapt our simulations for a multi-rate network, in which flows FTP-1 and FTP-2 transmit at 11 Mbps, flows FTP-3 and FTP-4 use 5.5 Mbps, and flows FTP-5 and FTP-6 use 2 Mbps. We also vary the packet size set for each flow for different simulations.

[084]   Figure 2 shows the per-flow throughput, as well as the overall throughput. It shows that our MWFS method achieves an aggregate throughput of 159% when compared with the prior art WFS method in the multi-rate wireless network. The throughput is similar to WFS when all packet flows use the same transmission rate. Significant throughput improvement is also achieved on a per-flow basis, particularly for flows that

use higher transmission rates. For example, throughput increases by 550% for flows FTP-1 and FTPO-2, and increases by 270% for flows FTO-3 and FTP-4 comparing with the case where each flow uses the base transmission rate of 2 Mbps.

[085]   Our MWFS allocates time periods 'fairly' among all packet flows, such that a high-rate packet flow is not penalized in the presence of low-rate packet flows. The prior art would have to give lower rate flows longer periods of time to achieve equal throughput, thus, severely degrading the performance of high-rate packet flows. Our MWFS method attains a significant increased throughput for both overall packet flow and on a per-flows basis.

[086]   Delay in Error-Free Channels

[087]   We further describe packet-delay performance of the MWFS method under error-free channel conditions. The arrival processes of the six flows or 'sources' are as follows: two MMPP flows, in which the modulated process is a continuous Markov chain with two states ON and OFF. The transition probability from the ON to OFF is 0.9, and the probability from state OFF to On is 0.1. Packets arrive according to a Poisson process at rate of 3.0. There are also two Poisson flows with packet arrival rate at 2.0. The remaining two flows are CBR sources with a packet generated every 1 ms. To measure the delay performance of the MWFS method with respect to different flows, we let flows MMPP-1, Poisson-1, CBR-1 transmit at the base transmission rate 2 Mbps and other flows transmit at 11 Mbps. The metrics used in the evaluation include an average delay $Delay^{avg}$ of

successfully transmitted packets, a maximum delay $Delay^{max}$, a standard delay deviation $Delay^{dev}$, and an average buffer size $Buffer^{avg}$ for each flow.

[088]   Figure 3 shows the results of this simulation. The delay experienced by the 11 Mbps flows is less than the 2Mbps for both methods. However, our MWFS method achieves substantially smaller delay time and a smaller buffer size for high-rate flows compared with the prior art WFS method. With the MWFS method, delays experienced by high-rate flows do not affect lower-rate flows. This demonstrates the capability of the MWFS method to provide differentiated services for high-rate and low-rate flows.

[089]   Throughput in Error-Prone Channels

[090]   In this simulation, we describe the effectiveness of the compensation model of the MWFS method for an error-prone channel. Ideally, the MWFS method preserves temporal fairness to achieve higher throughput and ensures a graceful compensation for lagging flows. For simplicity of presentation, we consider a scenario with four FTP flows. Two flows (FTP-3,4) use a base transmission rate of 2.0 Mbps and the other two flows (FTP-1,2) transmit at a higher rate of 11 Mbps. The channel error patterns of these flows are location-dependent and modeled by a two-state discrete Markov chain. By setting $p_e + p_g = 0.1$,we emulate highly correlated channel conditions for each flow.

[091]   Figure 4 shows the throughput results for the prior art WFS and the MWFS method according to the invention, where packet error rate varies from 0% up to 30%. As the packet error rate increases to 20%, the

throughputs for both scheduling methods decreases because the probability that all flows are simultaneously error-prone increases. However, the overall throughput of the MWFS method is substantially better than the WFS method.

[092]   Figure 5 shows the temporal fairness and effectiveness of our compensation model. There, we measure normalized time periods acquired by each flow. When the channel is error-free, each flow obtains a substantially equal amount of time 1.0. As the packet error rate increases, the length of the time periods obtained by flows decrease, which subsequently leads to an overall throughput reduction, as shown in Figure 4. However, the long-term temporal shares of all flows are still preserved, and are substantially equal for all packet flows. This shows that our compensation model is effective. Even when packet error rate increases to 20%, the throughput and time share gained by each flow reduce only slightly. This demonstrates that the MWFS method is still able to shield errors from flows and maintain a good overall throughput.

[093]   Packet Delay in Error-Prone Channel

[094]   We further describe the impact of an error-prone channel on packet delay. In this simulation setting, flows 1 and 2 are MMPP sources with packets arriving at the rate 1.0, and flows 3 and 4 are Poisson sources with packet arrival rate of 0.5. We also use two CBR flows with rate 1.0 to emulate the background traffic. The transmission rates for MMPP-1, Poisson-1 and CBR-1 are set as 2 Mbps, and the other three flows use a transmission rate of 11 Mbps.

22

[095]   Figures 5 and 6 show the packet delays for the MMPP and Poisson

flows, respectively. These Figures show that the delay experienced by each

flow increases as the channel error rate increases. Because the WFS method

maintains throughput fairness, the delays experienced by high-rate flows and

low-rate flows are approximately the same. However, delay performance is

different for low- and high-rate flows in the MWFS, because a temporal fair

share is used. Specifically, high-rate flows, MMPP-2 and Poisson-2,

experience noticeably less delay than the comparable low-rate flows.

However, low-rate flows have a delay performance comparable to the WFS.

This shows that the MWFS method is able to provide a certain degree of

flow separation among high-rate and low-rate flows, such that high-rate

flows are not penalized, or worse yet, paralyzed by low-rate packet flows.

[096]   Effect of the Invention

[097]   The invention provides a packet scheduling method that utilizes

the physical-layer multi-rate capability for WLANs based on the IEEE

802.11 standard. The models used by the method ensure fair temporal shares

for all terminals over extended periods of time. By taking advantage of

error-free channels, the method achieves significant throughput gains when

compared to prior art schedulers. Although the method is designed for

packet-switched WLAN based on IEEE 802.11b/a/g networks operating in

infrastructure mode, the method can also be used the IEEE 802.11e

framework by integrating the scheduler into a hybrid coordinator usually

collocated with the AP. The method is independent of the processes used to

estimate the channel estimation and transmission rates.

23

[098]   Although the invention has been described by way of examples of preferred embodiments, it is to be understood that various other adaptations and modifications may be made within the spirit and scope of the invention. Therefore, it is the object of the appended claims to cover all such variations and modifications as come within the true spirit and scope of the invention.